Rebuttal Report

Review of Principal Components Analysis of Data and Review of Inferences about Presence of Biomarkers in the Population of Animals from the Illinois River Watershed

Prepared for:
Tyson Foods, Inc.
Tyson Poultry, Inc.
Tyson Chicken, Inc.
Cobb-Vantress, Inc.
Cal-Maine Foods, Inc.
Cal-Maine Farms, Inc.
Cargill, Inc.
Cargill Turkey Production, LLC
George's, Inc.
George's Farms, Inc.
Peterson Farms, Inc.
Simmons Foods, Inc.
Willow Brook Farms, Inc.

Prepared by:
Charles D. Cowan, Ph.D.
Analytic Focus LLC
4939 De Zavala Road, Suite 105
San Antonio, TX 78249

November 26, 2008

Charles D. Cowan, Ph.D.

PERSONAL SUMMARY

1. My name is Charles Cowan. I reside in San Antonio, TX. I was retained by the defendants to provide an opinion regarding the use of principal components analysis by Dr. Olsen for this litigation and the statistical reliability and value of sampling used both by Dr. Olsen and Dr. Harwood. I have personal knowledge of the matters contained in this report.

Education and Experience

2. My background covers 30 years of research and study in the areas of statistics, economics, and their application to business problems. I am Managing Partner of Analytic Focus LLC, a company headquartered in San Antonio, TX and with offices in Birmingham, Alabama and Washington, DC. A portion of our work is conducting research for legal matters, including providing litigation support and expert witness services when requested. Some of our work focuses on measurement and mitigation of risk for financial intermediaries. The final area of our practice is in support of Federal and State agencies needing economic and financial analysis to pursue their missions. Prior to starting Analytic Focus LLC I served as Chief Statistician for the Federal Deposit Insurance Corporation. I was also a Director for Price Waterhouse where I headed the Financial Services Group in the Quantitative Methods Division. I served for 12 years at the U.S. Bureau of the Census where I was responsible for the evaluation of the Decennial Census and held the title of Chief of the Survey Design Branch.

3. I am currently an adjunct professor in the School of Public Health at the University of Alabama – Birmingham (UAB) and previously served as a professor in the Business School at UAB, as a visiting research professor at the University of Illinois, and in other academic and professional positions.

4. A listing of my qualifications as an expert in this case are presented in Appendix 1. My complete resume and a listing of all my publications are presented in Appendix 2. A listing of past cases in which I have been deposed or presented testimony at trial is presented in Appendix 3.

Scope of Assignment & Compensation

5. I was asked to consider the claims made by the plaintiffs in the above referenced case and to offer an opinion on issues pertaining to their claims. This report considers both issues.

Personnel	Fees per Hour
Charles Cowan, Ph.D.	$425
Senior Financial Analyst	$395
Senior Research Associate	$295
Programmer	$225
Research Analyst	$125

For expert representation, depositions and testimony, our hourly rate is $525. Out-of-pocket expenses, including travel, are billed separately and are in addition to the hourly fees.

EXPERT REPORT OF VALERIE HARWOOD

115. Dr. Harwood claims to have discovered a bacterium that carries a marker that uniquely identifies poultry litter – the poultry litter biomarker (PLB)[19]. Using this marker, she claims it is possible "to detect and quantify the amount of poultry-specific contamination in environmental samples, including soil, edge of field, surface water, and ground water samples collected in the IRW."[20]

116. In her report, Dr. Harwood explains the importance of determining the sensitivity and selectivity of a biomarker. "Sensitivity (the frequency of positive results when the contaminating source is present) and specificity (the frequency of negative results when the contaminating source is absent) are among the most important attributes of a useful MST test."[21] Failure to establish either of these characteristic renders a test useless, since high error rates mean that use of the biomarker is not reliable.

117. Dr. Harwood does not establish that the marker she claims to identify poultry litter has the requisite sensitivity or selectivity. Her sample is too small to measure any of the standards she establishes for either sensitivity or selectivity. Further, the cavalier sampling procedure used by the state to collect materials for testing is so flawed as to render it useless. This report addresses the statistical problems with Dr. Harwood's work and shows that none of her conclusions can be supported.

[19] CDM Report, page 6-31
[20] Harwood Report, paragraph 43, page 17
[21] Harwood Report, paragraph 42, page 17

General Principals of Sampling

118. There are four common sense issues to consider in drawing a sample. The first is determining what is to be measured. The second is the precision required for the estimate. The third is whether the sample is "representative". The term representative simply means that a sample selected from a population can be used to draw inferences about that population. The fourth is how to sample the population.

What to Measure

119. Determining what is to be measured is harder than it may seem. In Dr. Harwood's case, she wanted to know two statistics – how frequently does the marker show up in poultry, and how frequently is it absent in other animals. The latter is much more difficult to measure, since you are looking for something that may or may not be there.

120. However, this is a very common problem in statistics. For example, in biostatistics and epidemiology health researchers frequently face the problem of detection. The problem of detection is that, when an event occurs, we need to know about the event. The Centers for Disease Control (CDC) monitors hospitals for outbreaks of unusual diseases or patterns of diseases. The EPA and NASA are working jointly to combine satellite data with EPA tower data to detect particulate air pollution. In quality control, the object of the collection of data is to find that flaws do not exist in the product or process. A similar process is used in accounting and audit. In all of these cases, a reasonable sample size is necessary to ascertain that the error rate is below a certain threshold.

121. A simple example follows. Suppose I want to know how many people in the United States have a birthday in January. Assume that all months have equal numbers of days and that birthdays are spread uniformly throughout the year, so the chance of being born in January is one-twelfth. If I took a random sample of size 12, I'd expect to get one January birthday on average, but in fact I only have a 38% chance of getting one person with a January birthday. I have a 62% chance of getting zero or two. So if I take a sample of size 12 (a sample that would be large relative to what Dr. Harwood uses), I could extrapolate that result to the U.S. population, but I'd know very little about the number of people who have a birthday in January. In fact, I could easily conclude that NO ONE in the population has a birthday in January – but I'd be wrong. Now suppose that I want to know the number of people in the U.S. who have a birthday on January 1 in a non-leap year. The chance with the same small sample size of getting someone with a **rarer** event is much lower – in this case there is a 97% chance of getting no one with the January 1 birthday. But this does not mean that no one in the U.S. has a birthday on January 1.

122. But that's not the problem that Dr. Harwood is attempting to address. In the example above, I start with the conditions where **I know the probability** (because there are 12 months and a uniform distribution of births in my example). In the above example, I only want to know what the possible outcomes are. BUT Dr. Harwood doesn't know the relative frequency of the biomarker in poultry or more importantly in other species. For the example I gave above, where we know the probability is 1 in 12 (.083), we can determine how likely it is to get **no** positive outcomes. In the case where we don't know what the frequency of the biomarker is, we have to estimate the frequency from what we don't see. If we sample 12 animals using independent sampling techniques and observe no animals carrying the biomarker, the estimate for the proportion of animals with the biomarker is as high as 32% with a one percent chance the

number is actually higher than that. It could be 32% of animals in the total population carry the biomarker. It could be 21%, it could be 5% - we have no way of knowing.

123. Note that, Dr. Harwood doesn't have samples as large as 12 – hers are on the order of two or three. That means that the probability that animals in the population carry the biomarker can be as high as 78%, even when she doesn't observe an animal with the biomarker.

How Precise?

124. In the examples given above, the discussion focused on one measurement – how many animals with a marker. Dr. Harwood compounds the difficulty of the problem by considering whether the rates she studies differ by whether they are in the IRW or outside.

125. For comparisons between groups, there is sampling variation associated with each group separately. There is a certain amount of sampling variation associated with samples from inside the IRW. There is also sampling variation associated with samples from outside the IRW. To compare the results between the groups, the amount of sampling variation would be expected to at least double if the sample sizes are the same, or more if the sample sizes are different.

126. From Dr. Harwood's Table 2 [22] , she shows samples of size 5 for beef cattle inside and 5 for beef cattle outside the IRW. With a sample of size 5, the best she could determine is that with no samples showing the biomarker, the prevalence of the biomarker is less than 60%. In other words, as many as 60% of the cattle in either group could have the biomarker and she

[22] Harwood Report, page 24

could still get a sample of five cattle with no biomarker. As for comparing the rates between two groups of cattle, each with five in the sample, it would be impossible to make any determination of the differences between two groups.

127. In the same table, she shows that she has only one swine sampled inside and one sampled outside. There is **NO** statistical test that would allow any comparison of a sample of one with a sample of one; it is impossible to calculate any variability around the sample estimate. For a sample of size one, the outcome is either 0% (no marker) or 100% (marker found) – not very informative. The reliability of an estimate for a characteristic in the population is based on the variability measured in the sample. With a sample of size one, the variability is infinite, meaning there is no reliability. In other words, by taking a sample of one swine, Dr. Harwood knows nothing about whether any other swine in the population of swine carry the marker.

128. In fact, the same can be said for all sample sizes in this table, all of which are of size 5 or less. They are all so small that no inferences can be drawn from any of these samples about the presence or absence of the biomarker. Further, it would be impossible in most cases to even compare inside the IRW to outside the IRW since the sample sizes are too small to permit the calculation – it isn't even possible to do the calculation since the calculation in some cases would involve dividing by zero, a mathematical impossibility.

129. Finally, the other issue of concern is the precision Dr. Harwood says is required for her results to be acceptable. From her first deposition, Dr. Harwood says that "generally, in PCR, your error rate should be 5 percent or less".[23]

130. To calculate an error rate, one has to set up a table of the form:

	Truth	
Test	Positive	Negative
Positive	a	b
Negative	c	d

where the entries a, b, c, and d are counts of outcomes.

131. To measure a number like five percent, either b+c has to be very low relative to a+d (to ensure that the number is much lower than five percent), or the sample size has to be large enough to make a determination of the actual proportion. Consider that for a sample of size 31 (the number in Harwood's Table 2), the only proportions that can be estimated are 0/31, 1/31, 2/31, and so on, or 0%, 3.2%, 6.4%, 9.6%, and so on. In other words, if there is more than one error, the test fails the standard set by Dr. Harwood with this sample, and that requires combining all samples from cattle, swine, ducks, geese, and humans and the assumption that the error rates are the same for all groups. If they differ between groups, the problem is compounded since there is another level of variability in the outcomes to control.

132. To summarize, Dr. Harwood cannot achieve her own standard:

- the sample size is too small to even measure a number like 5%,
- with multiple types of animals tested, the measurement is confounded,

[23] Harwood Deposition 1, page 266, lines 10 and 11

- she did not perform these tests even though she clearly stated they were necessary for the acceptance of the biomarker.

Representativeness

133. To ensure representativeness of a sample, there needs to be either a sampling procedure in place that catalogs where the population is to be found and how it will be sampled, or a model explaining the movement of the population and how it will be captured. For example, in a clinical trials setting, the sampling requires an assumption that the population is coming to hospitals and will be captured at random, and the real randomness in a clinical trial comes from the administration of a treatment or control to each person coming to the hospital.

134. In a experimental design setting, especially this one in the IRW, the population doesn't come to the researcher, so the researcher has to go to the population. There need to be standards as to how the sample is selected, the likelihood of selection of a unit in the sample, and methods to allow extrapolation of the results of the sample to the population.

135. Dr. Harwood has none of this in place. She doesn't know where the population is (e.g. dairy cows), how many there are, how they were selected, nor how the populations interact with one another. This latter point relates to the fact that many of the animals cohabitate in areas, and so are exposed to similar influences. Testing one group may be correlated with testing of another group, but Dr. Harwood wouldn't know this is the case since there is no sampling frame or documentation that describes how the sample was selected from the population, or even what the population is.

136. This point is made over and above the concerns expressed above about the inadequacy of the sample sizes. Even with large sample sizes, the sample can be nonrepresentative (and biased) and so not very useful for analysis. In this case, no effort of any type was made to ensure that the sample represented a population.

How the Samples Were Selected

137. The final point has to do with the method for sampling. Much of the "sampling" done in this case was cluster sampling. The analyses conducted by Dr. Harwood and all of the statistics she presents all rely on an assumption of simple random sampling. This simply didn't happen.

138. A cluster sample is where multiple observations are taken from the same location, typically to reduce the cost of collection. The problem is that, units that are in the same location are more likely to be similar than units from different locations. This means that the variability of the observations increases, because there is the variability due to individual variation, and additionally the variation associated with the clusters being sampled.

139. This is a common phenomenon in surveys – a necessary evil to reduce the cost of data collection. In a survey of a human population about income, people who live on the same block are more likely to have similar incomes than those on different blocks. It is common to sample four households per sampled block in a survey to measure income or unemployment, but once the interviewer has spoken to the first household, there is less useful information from the other three households on the same block relative to households on other blocks.

140. The same is true for Dr. Harwood's data collection. Sampling multiple cow pies within the same field leads to two problems. One is that the cows milling about together in the same

field are likely to be much more similar to one another than cows in different fields (same feed, same water supply, etc.). The second problem is even worse – what's to keep the researcher from collecting multiple cow pies from the same cow? It's not as if they are distinguishable on examination – unless you are following each cow.

141. Because of this, Dr. Harwood's data is even less "precise" than she would expect, since the sampling methods used lead to significant increases in the variability of the outcomes. For this and for all the preceding reasons, Dr. Harwood's findings are meaningless.